Digital Impact: New Rating Cultures Challenge Academic Science

Franzen, Martina

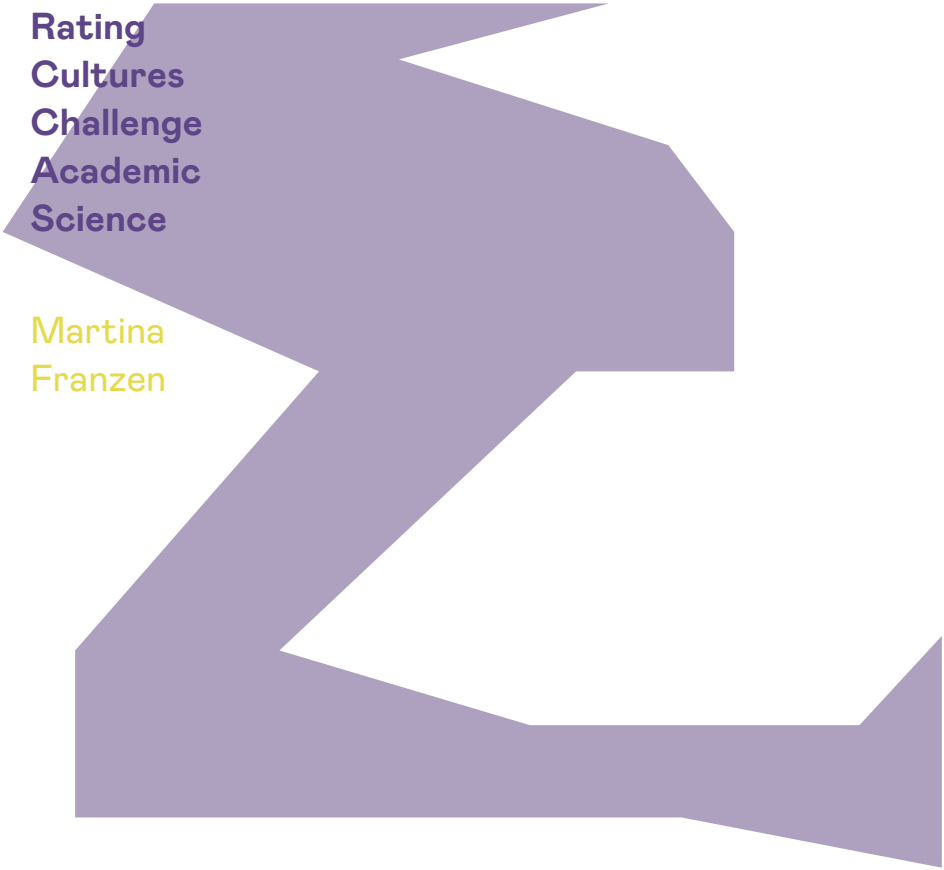
Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Franzen, Martina

Digital Impact: New Rating Cultures Challenge Academic Science

Digital Impact: New Rating Cultures Challenge Academic Science

Martina Franzen



"How satisfied are you with our service? Did the product meet your expectations?" Today, nearly every transaction in the digital economy comes with a request to give evaluative feedback. User or customer views are collected to optimize products or to improve marketing. Aside from ratings on various scales, customers are often asked to give written feedback in the form of reviews, which then may be rated by other customers to create a hierarchy out of the multiplicity of reviews. "Did you find this review helpful?" Positive responses boost the reviewer's calculated reputational rating, helping them climb the ranks of top reviewers, which serves as an incentive to write even more reviews.

One episode of the Netflix series *Black Mirror* unfolds a detailed scenario in which the ubiquitous John Doe rating serves as a new form of social control. The show paints the grim picture of a society based on a system of mutual ratings, in which the individual's only concern in every social interaction is getting as many likes as possible to increase their own reputational score. Ratings are given for every encounter or service. The rating is done by both parties, in real time and available online for all to see. The protagonist, Lacie, provides a great illustration of how the person-centred score governs individual behavior as soon as reaching a high numerical rating is not only motivated by one's narcissistic needs but necessary to obtain a certain socioeconomic status. For ambitious Lacie, a seemingly small difference between 4.2 and 4.5 on the five-point rating scale becomes an insurmountable obstacle on her path towards upward social mobility.

Like the Lacie character in *Black Mirror*, who hires a consultant to give strategic advice on how to raise her reputational rating as quickly as possible (to move into a more luxurious residence), researchers too can take advantage of numerous pieces of advice—some sincere, some cynical—to maximize their own impact rating. One article on the networking site Academia.edu, alluringly entitled "How to Increase Your Papers' Citation and H Index," (Gola, n.d.) has already garnered some 50,000 views. The author's ironic strategic recommendation: drastically increase the number of self-citations to attract

the necessary attention to your work. This short piece, which reveals the ethically questionable publishing practices of an Indonesian physics professor and presumably those of the author's colleagues, illustrates a problem of indicator-based performance assessments heavily discussed in academia: The focus on usage statistics in evaluative practice triggers gaming activities that undermine the meritocratic principle of equal performance assessment based on scientific quality criteria, possibly leading to an erosion of trust.

For quite some time, scientists in particular have been arguing over the extent to which qualitative characteristics may be translated into quantitative measures in a meaningful way. In addition to the classic instrument for qualitative assessments (i.e. the peer review system), the introduction of New Public Management at higher education institutions has added quantitative indicators, for instance when it comes to allocating grant money.

As evaluation research has shown a while ago, any kind of output control using quantitative indicators is accompanied by a neglect of content (Osterloh 2010). One-dimensional indicators may cause trade-offs in the system (Espeland and Sauder 2007). Marshall W. Meyer and Vipin Gupta (1994) speak of a "performance paradox" if indicators can no longer be used to distinguish strong performance from poor performance. When it comes to citation-based indicators in science such as the h-index or the Journal Impact Factor, "gaming the system" takes place on various levels: It concerns authors, editors, and publishers. The means to sanction ethically questionable publishing or citation practices are limited, as we all know.

Digitalization adds a new dimension to the focus on impact rates in science: The neologism *altmetrics* was coined to refer to methods for measuring a wide spectrum of web reactions to publications. The concept is fueled by the impetus to democratize science by creating an open and fairer system of performance assessment. That, in any case, was the thrust of the 2010 altmetrics manifesto (Priem et al. 2010), which served as the discursive cornerstone for further socio-technical development.

Altmetrics incorporate the full spectrum of research outputs such as journal articles, books, datasets, blog posts, and slide sets, as well as the multiple ways in which these outputs are used below the citation level (e.g. bookmarks, downloads, views). Unlike journal- or author-level metrics, altmetrics are an article- or rather an output-level rating tool. Instead of considering only the citation statistics of a set of source journals, such as those listed in Web of Science or Scopus, web-based measures refer to a repertory of sources that can be expanded to include all kinds of sources. If we take the service provider Altmetric.com as an example, the range of defined sources for the automatic measuring of impact includes social networks such as Facebook, microblogging services such as Twitter, video platforms such as YouTube, as well as international and national media outlets. However, Altmetric.com—a portfolio company of Digital Science, a subsidiary of MacMillan Publishers Ltd.—is best known for its attention score. Based on an undisclosed algorithm, the Altmetric score is displayed in the form of so-called badges. One of the most popular badges is the Altmetric donut: a ring whose coloring offers information about the type of achieved impact, that is, about sources (blue for Twitter, red for newspapers, and so forth). A nice technical gadget, one might think, but irrelevant for science. The proponents of altmetrics, who are found in parts of academia, the IT sector, libraries, and scientific publishing houses, think differently. They want altmetrics to become the catalyst in revamping the academic reputation system. But how are Tweets or Facebook likes supposed to tell us anything about scientific quality or relevance?

Even as this decisive question in terms of methodology remains unanswered, the comprehensive implementation of altmetrics tools in digital publication infrastructures continues. Large international publishers such as Elsevier, Wiley, or Springer, as well as the top journals *Science* and *Nature* have already integrated them into their portfolio. The social network ResearchGate also uses altmetrics based on the collected publication data and the personalized usage statistics—the one difference being that ResearchGate additionally provides a person-centred score. This score puts researchers in relation to one another. Transparency is created by showing ResearchGate members

and all their readers the exact calculation of the percentile into which the individual score falls. The score is cumulative but subject to minor and sometimes confusing ups and even downs. Users receive weekly statistics detailing the usage of their own and other people's contributions. To keep users motivated, ResearchGate transplanted a classic feature of the gaming sector into scientific communications: announcing users' entry into a new level, based here on achieving a certain threshold of citations or clicks or a top position in the institutional ranking. As in digital gaming environments, ResearchGate too provides users with tips on how to raise their individual score, such as: "Boost your stats by adding more research."

By means of such incentive systems, the digital platforms ResearchGate and Academia.edu gradually collect more and more data of all kinds. Big data, therefore, is the foundation of their business model, the outlines of which have so far been blurry at best. In 2016, Academia.edu, the US counterpart of the German start-up ResearchGate, introduced a premium account option for an annual membership fee of 99 US dollars. As far as content is concerned, there is still hardly any difference between the premium feature and the freemium account—the only difference is that premium members have access to a detailed overview of how each of their contributions is used, including user or reader characteristics, listed by person or aggregated by institutions, countries, and so forth.

Formerly, digital platforms focused on rating scientists as authors; now, scientists are also measured as readers with regard to their individual usage patterns. It seems doubtful that paywalls can be established in scientific communications, given the strength of the open science movement. But the offer to learn more about who reads your publications meets the genuine needs of researchers, who—unlike literary authors—cannot turn to book sales to get an idea of their publications' reach. Whereas traditional citation measures only showed the tip of the iceberg, altmetrics now show the full scope of how research output is used beyond formal citations in scientific journals, making that usage the basis on which scientists are rated. This approach satisfies the narcissistic needs of researchers and possibly offers extra

informational value for institutional research evaluation. The key question, however—what do altmetrics actually measure—remains unanswered.

The dominant research approach in bibliometrics (i.e. conducting empirical studies comparing citation rates and altmetrics of all kinds) does not help much in this case. To be sure, citation may be theoretically conceived of as a form of social recognition of scientific achievement. But trying to identify differences in scientific quality based on the sheer number of citations leads to a short circuit between impact (i.e. popularity) and quality. Based on my own work on the medialization of science (Franzen 2011, 2015), I propose a different assumption: First and foremost, altmetrics—like citation rates—signal popularity. High impact rates may in fact coincide with scientific quality, but they may also result from news factors such as entertainment, scandals, or celebrity. The explanatory power of altmetrics (and citation rates) may thus be reduced primarily to measuring marketing success. Marketing success—in the sense of achieving high impact rates—can indeed be an indicator of special scientific quality, but the political sensitivity or currency of an issue, the prominence of the author, or simply well-placed advertising are equally conducive to high impact.

One good example to illustrate the argument that scientific ratings may conflict with news ratings when it comes to measuring impact is the annual ranking of the top 100 articles per Altmetric score. In 2016, the number one article appeared in the prestigious *Journal of the American Medical Association*. Its author, however, is not a medical researcher, as one might expect, but the then President of the United States of America, Barack Obama himself, writing about US healthcare reform (Altmetric, n.d.). It is obvious that the honor of getting the highest Altmetric score has little to do with criteria of scientific relevance. Against this background, it is even more surprising that altmetrics have hardly been questioned in the scientific community.

Paul Wouters and Rodrigo Costas (2012) have referred to the altmetrics concept as a "narcissistic technology." This presumably also explains its rapid rise. The question is: Will

it continue to enjoy this immense popularity once it mutates into an actual "monitoring technology"? Although altmetrics have not yet been officially introduced as an evaluation tool into institutional performance assessments, their implementation, for instance for measuring societal impact, seems only a matter of time. But any kind of performance assessment is bound to trigger a behavioral response and is not without consequences for the system. The kind of reactivity criticized as "gaming the system" may also be viewed as a successful adaptation to misguided indicators. The game is an old one: With the impact factor, gaming primarily involved the journals (via editorial choices and PR); with altmetrics, it is now the authors themselves who come into play. Their job is to engage in successful reputation management and to steadily boost their own click rates by advertising themselves on social media, rating other people's work, or communicating with just the right target groups. The pursuit of maximum reach, however, requires different means and is not a genuine goal of scientific work. Rather, it is a response to the conditions of the attention economy in the digital age – including all the possible consequences with regard to the quality of the produced knowledge in the overall process of knowledge formation.

References

Altmetric. n.d. “Top 100 Articles 2016.” Accessed January 1, 2017. https://www.altmetric.com/top100/2016/.

Brooker, Charlie. 2016. “Nosedive.” *Black Mirror*. Season 3. Episode 1. Los Gatos, CA: Netflix.

Dem, Gola. n.d. “How to Increase Your Papers’ Citation and H Index.” Accessed January 1, 2017. https://www.academia.edu/934257/How_to_increase_your_papers_citations_and_h_index_in_5_simple_steps.

Espeland, Wendy N., and Michael Sauder. 2007. “Rankings and Reactivity: How Public Measures Recreate Social Worlds.” In *American Journal of Sociology* 113 (1): 1–40.

Franzen, Martina. 2011. *Breaking News: Wissenschaftliche Zeitschriften im Kampf um Aufmerksamkeit*. Baden-Baden: Nomos.

Franzen, Martina. 2015. “Der Impact Faktor war gestern: Altmetrics und die Zukunft der Wissenschaft“. Themenheft: Der impact des impact factors. In *Soziale Welt* 66 (2): 225–42.

Meyer, Marshall W., and Vipin Gupta. 1994. “The Performance Paradox.” *In Research in Organizational Behavior* 16: 309–69.

Osterloh, Margrit. 2010. “Governance by Numbers: Does It Really Work in Research?“ In *Analyse und Kritik* 32 (2): 267–83.

Priem, Jason, Dario Taraborelli, Paul Groth, and Cameron Neylon. 2010. “Altmetrics: A manifesto.” Accessed September 27, 2017. http://altmetrics.org/manifesto/.

Wouters, Paul, and Rodrigo Costas. 2012. “Users, Narcissism and Control – Tracking the Impact of Scholarly Publications in the 21st Century“. Utrecht: SURFfoundation. Accessed January 1, 2017. http://research-acumen.eu/wp-content/uploads/Users-narcissism-and-control.pdf.

First published in German as “Digitale Resonanz: Neue Bewertungskulturen fordern die Wissenschaft heraus” in *WZB Mitteilungen*, Quarterly Journal of the WZB Berlin Social Science Center, 155 (March 2017): 30–33.